

ENTERED



1600

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/08/520,946A

DATE: 06/21/2002  
TIME: 14:46:47

Input Set : A:\Seqsub2.app  
Output Set: N:\CRF3\06212002\H520946A.raw

RECEIVED

JUL 08 2002

## SEQUENCE LISTING

TECH CENTER 1600/2900

## 4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: BROW, MARY ANN D.  
7 LYAMICHEV, VICTOR I.  
8 OLIVE, DAVID M.

10 (ii) TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF  
11 PATHOGENS

13 (iii) NUMBER OF SEQUENCES: 165

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: MEDLEN & CARROLL  
17 (B) STREET: 220 MONTGOMERY STREET, SUITE 2200  
18 (C) CITY: SAN FRANCISCO  
19 (D) STATE: CALIFORNIA  
20 (E) COUNTRY: UNITED STATES OF AMERICA  
21 (F) ZIP: 94104

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk  
25 (B) COMPUTER: IBM PC compatible  
26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/08/520,946A  
C--> 31 (B) FILING DATE: 30-Aug-1995  
32 (C) CLASSIFICATION:

34 (viii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: CARROLL, PETER G.  
36 (B) REGISTRATION NUMBER: 32,837  
37 (C) REFERENCE/DOCKET NUMBER: FORS-01756

39 (ix) TELECOMMUNICATION INFORMATION:

40 (A) TELEPHONE: (415) 705-8410  
41 (B) TELEFAX: (415) 397-8338

44 (2) INFORMATION FOR SEQ ID NO: 1:

46 (i) SEQUENCE CHARACTERISTICS:

47 (A) LENGTH: 2506 base pairs  
48 (B) TYPE: nucleic acid  
49 (C) STRANDEDNESS: double  
50 (D) TOPOLOGY: linear

52 (ii) MOLECULE TYPE: DNA (genomic)

56 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

58 ATGAGGGGGA TGCTGCCCT CTTTGAGCCC AAGGGCCGGG TCCTCCTGGT GGACGCCAC 60

60 CACCTGGCCT ACCGCACCTT CCACGCCCTG AAGGGCCTCA CCACCAGCCG GGGGGAGCCG 120

62 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCCTCAAGG CCCTCAAGGA GGACGGGGAC 180

64 GCGGTGATCG TGGTCTTGA CGCCAAGGCC CCCTCCTTCC GCCACGAGGC CTACGGGGGG 240

RECEIVED  
JUL 28 2003  
TECH CENTER 1600/2900

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/08/520,946A

DATE: 06/21/2002  
TIME: 14:46:47

Input Set : A:\Seqsub2.app  
Output Set: N:\CRF3\06212002\H520946A.raw

|     |             |            |             |            |             |             |      |
|-----|-------------|------------|-------------|------------|-------------|-------------|------|
| 66  | TACAAGGCCGG | GCCGGGCCCC | CACGCCGGAG  | GACTTTCCC  | GGCAACTCGC  | CCTCATCAAG  | 300  |
| 68  | GAGCTGGTGG  | ACCTCCTGGG | GCTGGCGCG   | CTCGAGGTCC | CGGGCTACGA  | GGCGGACGAC  | 360  |
| 70  | GTCCTGGCCA  | GCCTGGCCAA | GAAGGCCGAA  | AAGGAGGGCT | ACGAGGTCCG  | CATCCTCACC  | 420  |
| 72  | GCCGACAAAG  | ACCTTACCA  | GCTCCTTCC   | GACCGCATCC | ACGTCCCTCCA | CCCCGAGGGG  | 480  |
| 74  | TACCTCATCA  | CCCCGGCCTG | GCTTGGGAA   | AACTACGGCC | TGAGGCCGGA  | CCAGTGGGCC  | 540  |
| 76  | GACTACCGGG  | CCCTGACCGG | GGACGAGTCC  | GACAACCTTC | CCGGGGTCAA  | GGGCATCGGG  | 600  |
| 78  | GAGAACACGG  | CGAGGAAGCT | TCTGGAGGAG  | TGGGGGAGCC | TGGAAGCCCT  | CCTCAAGAAC  | 660  |
| 80  | CTGGACCGGC  | TGAAGCCGAG | AAGATCCTGG  | CCCACATGGA | CGATCTGAAG  | 720         |      |
| 82  | CTCTCCTGGG  | ACCTGGCCAA | GGTGCCACC   | GACCTGCCCC | TGGAGGTGGA  | CTTCGCCAAA  | 780  |
| 84  | AGGCAGGGAGC | CCGACCGGGA | GAGGCCCTAGG | GCCTTCTGG  | AGAGGCTTG   | GTTTGGCAGC  | 840  |
| 86  | CTCCTCCACG  | AGTTCCGCT  | TCTGGAAAGC  | CCCAAGGCC  | TGGAGGAGGC  | CCCCCTGGCCC | 900  |
| 88  | CCGCGGAAG   | GGGCCTTCGT | GGGCTTGTG   | CTTCCCAGCA | AGGAGCCCAT  | GTGGGCCGAT  | 960  |
| 90  | CTTCTGGCCC  | TGGCCGCCGC | CAGGGGGGGC  | CGGGTCCACC | GGGCCCGCGA  | GCCTTATAAA  | 1020 |
| 92  | GCCCTCAGGG  | ACCTGAAGGA | GGCGGGGGG   | CTTCTCGCCA | AAGACCTGAG  | CGTTCTGGCC  | 1080 |
| 94  | CTGAGGGAAAG | GCCTTGGCCT | CCGCGCCGGC  | GACGACCCCC | TGCTCCTCGC  | CTACCTCCTG  | 1140 |
| 96  | GACCTTCCA   | ACACCACCCC | CGAGGGGGTG  | GCCCAGCGCT | ACGGCGGGGA  | GTGGACGGAG  | 1200 |
| 98  | GAGGCAGGGGG | AGCGGGCCGC | CCTTCCGAG   | AGGCTCTTCG | CCAACCTGTG  | GGGGAGGCTT  | 1260 |
| 100 | GAGGGGGAGG  | AGAGGCTCT  | TTGGCTTAC   | CGGGAGGTGG | AGAGGCCCT   | TTCCGCTGTC  | 1320 |
| 102 | CTGGCCCACA  | TGGAGGCCAC | GGGGGTGCGC  | CTGGACGTGG | CCTATCTCAG  | GGCCTTGTCC  | 1380 |
| 104 | CTGGAGGTGG  | CCGAGGAGAT | CGCCCGCCCT  | GAGGCCGAGG | TCTTCCGCT   | GGCCGGCCAC  | 1440 |
| 106 | CCCTTCAACC  | TCAACTCCCG | GGACCACTG   | GAAAGGGTCC | TCTTGACGA   | GCTAGGGCTT  | 1500 |
| 108 | CCCGCCATCG  | GCAAGACGGA | GAAGACCGGC  | AAGCGCTCCA | CCAGCGCCGC  | CGTCTGGAG   | 1560 |
| 110 | GCCCTCCCG   | AGGCCACCC  | CATCGTGGAG  | AAGATCCTGC | AGTACCGGGA  | GCTCACCAAG  | 1620 |
| 112 | CTGAAGAGCA  | CCTACATTGA | CCCTTGGCG   | GACCTCATCC | ACCCCAAGGAC | GGCCGCCTC   | 1680 |
| 114 | CACACCGCT   | TCAACCAGAC | GGCCACGGCC  | ACGGCAGGC  | TAAGTAGCTC  | CGATCCCAAC  | 1740 |
| 116 | CTCCAGAAC   | TCCCCGTCCG | CACCCCGCTT  | GGGCAGAGGA | TCCGCCGGC   | CTTCATCGCC  | 1800 |
| 118 | GAGGAGGGGT  | GGCTATTGGT | GGCCCTGGAC  | TATAGCCAGA | TAGAGCTCAG  | GGTGTGGCC   | 1860 |
| 120 | CACCTCTCCG  | GCGACGAGAA | CCTGATCCGG  | GTCTTCCAGG | AGGGGCGGGG  | CATCCACACG  | 1920 |
| 122 | GAGACGCCA   | GCTGGATGTT | CGGCGTCCCC  | CGGGAGGCCG | TGGACCCCT   | GATGCCGCCG  | 1980 |
| 124 | GCGGCCAAGA  | CCATCAACTT | CGGGGTCCCT  | TACGGCATGT | CGGCCACCG   | CCTCTCCCAG  | 2040 |
| 126 | GAGCTAGCCA  | TCCCTTACGA | GGAGGCCAG   | GCCTCATG   | AGCGCTACTT  | TCAGAGCTTC  | 2100 |
| 128 | CCCAAGGTGC  | GGGCCTGGAT | TGAGAACACC  | CTGGAGGAGG | GCAGGAGGCG  | GGGGTACGTG  | 2160 |
| 130 | GAGACCTCT   | TCGGCCGCCG | CCGCTACCGT  | CCAGACCTAG | AGGCCCGGGT  | GAAGAGCGTG  | 2220 |
| 132 | CGGGAGGCCG  | CCGAGCGCAT | GGCCTTCAAC  | ATGCCCGTCC | AGGGCACCGC  | CGCCGACCTC  | 2280 |
| 134 | ATGAAGCTGG  | CTATGGTGAA | GCTCTTCCCC  | AGGCTGGAGG | AAATGGGGGC  | CAGGATGCTC  | 2340 |
| 136 | CTTCAGGTCC  | ACGACGAGCT | GGTCCTCGAG  | GCCCCAAAAG | AGAGGGCGGA  | GGCCGTGGCC  | 2400 |
| 138 | CGGCTGGCCA  | AGGAGGTCA  | GGAGGGGGTG  | TATCCCCTGG | CCGTGCCCT   | GGAGGTGGAG  | 2460 |
| 140 | GTGGGGATAG  | GGGAGGACTG | GCTCTCCGCC  | AAGGAGTGAT | ACCACC      |             | 2506 |

142 (2) INFORMATION FOR SEQ ID NO: 2:

144 (i) SEQUENCE CHARACTERISTICS:

145 (A) LENGTH: 2496 base pairs

146 (B) TYPE: nucleic acid

147 (C) STRANDEDNESS: double

148 (D) TOPOLOGY: linear

150 (ii) MOLECULE TYPE: DNA (genomic)

154 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

|     |            |            |           |            |           |            |    |
|-----|------------|------------|-----------|------------|-----------|------------|----|
| 156 | ATGGCGATGC | TTCCCCCTTT | TGAGCCAAA | GGCCGCGTGC | TCCTGGTGG | CGGCCACCAC | 60 |
|-----|------------|------------|-----------|------------|-----------|------------|----|

|     |            |            |            |            |            |            |     |
|-----|------------|------------|------------|------------|------------|------------|-----|
| 158 | CTGGCCTACC | GCACCTTCTT | TGCCCTCAAG | GGCCTCACCA | CCAGCCGCGG | CGAACCCGTT | 120 |
|-----|------------|------------|------------|------------|------------|------------|-----|

|     |            |            |           |           |            |            |     |
|-----|------------|------------|-----------|-----------|------------|------------|-----|
| 160 | CAGGCGGTCT | ACGGCTTCGC | CAAAGCCTC | CTCAAGGCC | TGAAGGAGGA | CGGGGACGTG | 180 |
|-----|------------|------------|-----------|-----------|------------|------------|-----|

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/08/520,946A

DATE: 06/21/2002  
TIME: 14:46:47

Input Set : A:\Seqsub2.app  
Output Set: N:\CRF3\06212002\H520946A.raw

|     |  |             |         |        |            |          |          |         |          |      |
|-----|--|-------------|---------|--------|------------|----------|----------|---------|----------|------|
| 162 | GTGGTGGTGG                               | TCTTGACGC   | CAAGGCC | CC     | TCCTTCCGCC | ACGAGGCC | TA       | CGAGGCC | TAC      | 240  |
| 164 | AAGGC                                    | GGGCC       | GGGCC   | CCC    | GAGGAC     | TTT      | CCCC     | GGC     | AGCTGGCC | 300  |
| 166 | TTGGTGGACC                               | TCCTAGGC    | TGTGCGG | CTG    | GAGGTT     | GGG      | CCTGAGGC | GGACGAC | GTG      | 360  |
| 168 | CTGGCCACCC                               | TGGCCAAGCG  | GGCGGAA | AAAG   | GAGGGT     | TACG     | AGGTGCG  | CAT     | CCTCA    | 420  |
| 170 | GACCGCGACC                               | TCTACCAGCT  | CCTTC   | GGGAG  | CGCAT      | CGGCC    | TCCTCC   | ACCC    | TGAGGGT  | 480  |
| 172 | CTGATCACCC                               | CGGCGTGGCT  | TTACG   | GAGAAG | TACGGC     | CTGC     | GCCC     | GGAGCA  | GTGGG    | 540  |
| 174 | TACCGGGCCC                               | TGGCGGGGGA  | CCCCT   | CGGAT  | AACAT      | CCCC     | GGGT     | GAAGGG  | CAT      | 600  |
| 176 | AAGACCGCCC                               | AGAGGCTCAT  | CCGCG   | GAGTGG | GGGAGC     | CTGG     | AAAAC    | CTT     | CCAGCAC  | 660  |
| 178 | GACCAGGTGA                               | AGCCTCCTT   | GGCGG   | GAGAAG | CTCCAG     | GGGG     | GCAT     | GGAGGC  | CCTG     | 720  |
| 180 | TCCCAGGAAGC                              | TTTCCCAGGT  | GCACACT | GAC    | CTG        | CCCCT    | AGGTG    | GACTT   | CGGGAGG  | 780  |
| 182 | CGCACACCCA                               | ACCTGGAGGG  | TCTG    | CGGGG  | CTG        | TTTGG    | GGT      | TGGAGTT | TGGAAGC  | 840  |
| 184 | CTCCACGAGT                               | TCGGCCTCCT  | GGAGGG  | GGCC   | AAGGCG     | GGAGG    | CCC      | CTGG    | CCCCCT   | 900  |
| 186 | CCGGAAGGGG                               | CTTTTTGGG   | CTTTT   | CCCTT  | TCCC       | GGCCC    | AGCC     | CATGTG  | GGCC     | 960  |
| 188 | CTGGCCCTGG                               | CTGGGGCGT   | GGAGGG  | GGCG   | CTCC       | ATCGGG   | CACA     | AGACCC  | CCTT     | 1020 |
| 190 | CTGAGGGACC                               | TTAAGGGGG   | GC      | GGGAA  | ATC        | CTGG     | GGAGT    | ACCTG   | GGCG     | 1080 |
| 192 | CGGGAGGGCC                               | TGGACCTCTT  | CCCAG   | AGGAC  | GACCC      | CATG     | TCCT     | GGCTA   | CCTT     | 1140 |
| 194 | CCCTCAAACA                               | CCACCCCTGA  | GGGGG   | TGGCC  | CGGC       | GTTAC    | GGGGG    | GAGTG   | GACG     | 1200 |
| 196 | GCGGGGGAGA                               | GGGGCCCTCCT | GGCCG   | GAGCG  | CTCT       | CCAGA    | CCCTAA   | AGGA    | GCGC     | 1260 |
| 198 | GGAGAAGAAC                               | GCCTGTTTG   | GCTT    | TACGAG | GAGG       | TGGAGA   | AGCG     | CCTTC   | CGGG     | 1320 |
| 200 | GCCCCGATGG                               | AGGCCACGGG  | GGT     | CCGG   | CTG        | GAC      | GTGG     | CCT     | ACCT     | 1380 |
| 202 | GAGGTGGAGG                               | CGGAGGTGCG  | CCAGC   | TGGAG  | GAGG       | AGGTCT   | TCC      | GCTGG   | CGGC     | 1440 |
| 204 | TTCAACCTCA                               | ACTCCC      | GCGA    | CA     | CCAG       | CTGGAG   | CGGG     | TGCTCT  | TTG      | 1500 |
| 206 | GCCATCGCA                                | AGACGGAGAA  | GAC     | GGGG   | GGAA       | CGCT     | CCACCA   | GCG     | CTGCC    | 1560 |
| 208 | CTGCGAGAGG                               | CCCACCCCAT  | CGT     | GGACCG | CG         | ATC      | CTG      | CAGT    | ACCG     | 1620 |
| 210 | AAGAACACCT                               | ACATAGACCC  | CCT     | GCCC   | GGCC       | CTGG     | TCC      | AC      | CCAAGAC  | 1680 |
| 212 | ACCGCTCTCA                               | ACCAGACGG   | CAC     | CGCC   | AC         | CG       | AGG      | GCTT    | CCAG     | 1740 |
| 214 | CAGAACATCC                               | CCGTGCGC    | CC      | CT     | GGG        | CAG      | GGC      | ATCC    | CG       | 1800 |
| 216 | GAGGGCTGGG                               | TGCTGGTGGT  | CTT     | GGACT  | AC         | CCAG     | ATTG     | AG      | CTGGG    | 1860 |
| 218 | CTCTCCGGGG                               | ACGAGAACCT  | GAT     | CCGGG  | TC         | TT       | CAGG     | AGG     | GGAGG    | 1920 |
| 220 | ACCGCCAGCT                               | GGATGTTCGG  | CGT     | TT     | CCCC       | GAAGGG   | GTAG     | AC      | CC       | 1980 |
| 222 | GCCAAGACCA                               | TCAACTTC    | GGT     | GCT    | CTAC       | GGC      | ATG      | TCC     | GCG      | 2040 |
| 224 | CTTTCATCC                                | CCTACGAGGA  | GGC     | GGT    | GGC        | TTC      | ATTG     | GAG     | CTT      | 2100 |
| 226 | AAGGTGGGG                                | CCTGGATTGA  | GGG     | GGAC   | CTC        | GAG      | GAGG     | GGCC    | GGGG     | 2160 |
| 228 | ACCCCTTCG                                | GCCGCC      | CTAT    | GTG    | CCC        | ACG      | CCGGG    | GTGAA   | GAGCG    | 2220 |
| 230 | GAGGC                                    | GGCG        | AGCG    | CAT    | GGC        | ATG      | TCC      | AGG     | CCG      | 2280 |
| 232 | AAGCTGGCCA                               | TGGTGG      | TTT     | CCC    | CCG        | CTT      | CAGGA    | TGGGG   | GCGAG    | 2340 |
| 234 | CAGGTGCACG                               | ACGAGCTGGT  | CCT     | CGAG   | GCC        | CCC      | AAAGG    | ACC     | GGGG     | 2400 |
| 236 | TTGGCCAAGG                               | AGGTCA      | GGG     | GGG    | TCTGG      | CCC      | CTGC     | AGG     | TG       | 2460 |
| 238 | GGCCTGGGG                                | AGGACTGGCT  | CTCC    | GGCCA  | AG         | GAG      | TAG      |         |          | 2496 |
| 240 | (2) INFORMATION FOR SEQ ID NO: 3:        |             |         |        |            |          |          |         |          |      |
| 242 | (i) SEQUENCE CHARACTERISTICS:            |             |         |        |            |          |          |         |          |      |
| 243 | (A) LENGTH: 2504 base pairs              |             |         |        |            |          |          |         |          |      |
| 244 | (B) TYPE: nucleic acid                   |             |         |        |            |          |          |         |          |      |
| 245 | (C) STRANDEDNESS: double                 |             |         |        |            |          |          |         |          |      |
| 246 | (D) TOPOLOGY: linear                     |             |         |        |            |          |          |         |          |      |
| 248 | (ii) MOLECULE TYPE: DNA (genomic)        |             |         |        |            |          |          |         |          |      |
| 252 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: |             |         |        |            |          |          |         |          |      |
| 254 | ATGGAGGC                                 | GA          | TGCT    | CCG    | CTT        | GAA      | CC       | GGAC    | GGCC     | 60   |
| 256 | CACCTGGC                                 | CT          | ACCG    | CAC    | CT         | GG       | CC       | AGC     | GGAC     | 120  |

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/08/520,946A

DATE: 06/21/2002  
TIME: 14:46:47

Input Set : A:\Seqsub2.app  
Output Set: N:\CRF3\06212002\H520946A.raw

|     |             |             |             |            |            |            |      |
|-----|-------------|-------------|-------------|------------|------------|------------|------|
| 258 | GTGCAGGCGG  | TCTACGGCTT  | CGCCAAGAGC  | CTCCTCAAGG | CCCTGAAGGA | GGACGGGTAC | 180  |
| 260 | AAGGCCGTCT  | TCGTGGTCTT  | TGACGCCAAG  | GCCCCCTCCT | TCCGCCACGA | GGCCTACGAG | 240  |
| 262 | GCCTACAAGG  | CGGGGAGGGC  | CCCGACCCCC  | GAGGACTTCC | CCCGGCAGCT | CGCCCTCATC | 300  |
| 264 | AAGGAGCTGG  | TGGACCTCCT  | GGGGTTTACC  | CGCCTCGAGG | TCCCCGGCTA | CGAGGCGGAC | 360  |
| 266 | GACGTTCTCG  | CCACCCCTGGC | CAAGAAGGCG  | GAAAAGGAGG | GGTACGAGGT | GCGCATCCTC | 420  |
| 268 | ACCGCCGACC  | GCGACCTCTA  | CCAACTCGTC  | TCCGACCGCG | TCGCCGTCT  | CCACCCCGAG | 480  |
| 270 | GGCCACCTCA  | TCACCCCGGA  | GTGGCTTGG   | GAGAAGTACG | GCCTCAGGCC | GGAGCAGTGG | 540  |
| 272 | GTGGACTTCC  | GCGCCCTCGT  | GGGGGACCCC  | TCCGACAACC | TCCCCGGGGT | CAAGGGCATC | 600  |
| 274 | GGGGAGAAGA  | CCGCCCTCAA  | GCTCCTCAAG  | GAGTGGGGAA | GCCTGGAAA  | CCTCCTCAAG | 660  |
| 276 | AACCTGGACC  | GGGTAAAGCC  | AGAAAACGTC  | CGGGAGAAGA | TCAAGGCCA  | CCTGGAAGAC | 720  |
| 278 | CTCAGGCTCT  | CCTTGGAGCT  | CTCCCGGGTG  | CGCACCGACC | TCCCCCTGGA | GGTGGACCTC | 780  |
| 280 | GCCCAGGGGC  | GGGAGGCCGA  | CGGGGAGGGG  | CTTAGGGCCT | TCCTGGAGAG | GCTGGAGTTC | 840  |
| 282 | GGCAGCCTCC  | TCCACGAGTT  | CGGCCCTCTG  | GAGGGCCCCG | CCCCCCTGGA | GGAGGCCCCC | 900  |
| 284 | TGGCCCCCGC  | CGGAAGGGGC  | CTTCGTGGGC  | TTCTGCCTCT | CCCGCCCCGA | GCCCATGTGG | 960  |
| 286 | GCGGAGCTTA  | AAGCCCTGGC  | CGCCTGCAGG  | GACGGCCGGG | TGCACCGGGC | AGCAGACCCC | 1020 |
| 288 | TTGGCGGGGC  | TAAAGGACCT  | CAAGGAGGTC  | CGGGGCCTCC | TCGCCAAGGA | CCTGCCGTTC | 1080 |
| 290 | TTGGCCTCGA  | GGGAGGGGCT  | AGACCTCGTG  | CCCGGGGACG | ACCCCATGCT | CCTGCCCTAC | 1140 |
| 292 | CTCCTGGACC  | CCTCCAACAC  | CACCCCCGAG  | GGGGTGGCAG | GGCGCTACGG | GGGGGAGTGG | 1200 |
| 294 | ACGGAGGACG  | CGGCCAACCG  | GGCCCTCCTC  | TCGGAGAGGC | TCCATCGGAA | CCTCCTTAAG | 1260 |
| 296 | CGCCTCGAGG  | GGGAGGAGAA  | GCTCCTTGG   | CTCTACCACG | AGGTGGAAA  | GCCCTCTCC  | 1320 |
| 298 | CGGGTCCCTGG | CCACATGGA   | GGCCACCGGG  | GTACGGCTGG | ACGTGGCTTA | CCTTCAGGCC | 1380 |
| 300 | CTTCCCCCTGG | AGCTGCGGA   | GGAGATCCGC  | CGCCTCGAGG | AGGAGGTCTT | CCGCTTGGCG | 1440 |
| 302 | GGCCACCCCT  | TCAACCTCAA  | CTCCCGGGAC  | CAGCTGGAAA | GGGTGCTCTT | TGACGAGCTT | 1500 |
| 304 | AGGCTTCCCG  | CCTTGGGAA   | GACGCAAAG   | ACAGGCAAGC | GCTCCACCA  | CGCCGCGGTG | 1560 |
| 306 | CTGGAGGCC   | TACGGGAGGC  | CCACCCATC   | GTGGAGAAGA | TCCTCCAGCA | CGGGGAGCTC | 1620 |
| 308 | ACCAAGCTCA  | AGAACACCTA  | CGTGGACCCC  | CTCCCAAGCC | TCGTCCACCC | GAGGACGGGC | 1680 |
| 310 | CGCCTCCACA  | CCCGCTTCAA  | CCAGACGGCC  | ACGGCCACGG | GGAGGCTTAG | TAGCTCCGAC | 1740 |
| 312 | CCCAACCTGC  | AGAACATCCC  | CGTCCGCACC  | CCCTTGGGCC | AGAGGATCCG | CGGGCCTTC  | 1800 |
| 314 | GTGGCCGAGG  | CGGGTTGGGC  | GTTGGTGGCC  | CTGGACTATA | GCCAGATAGA | GCTCCCGTTC | 1860 |
| 316 | CTCGCCCACC  | TCTCCGGGGA  | CGAAAAACCTG | ATCAGGGTCT | TCCAGGAGGG | GAAGGACATC | 1920 |
| 318 | CACACCCAGA  | CCGCAAGCTG  | GATGTTCCGC  | GTCCCCCGG  | AGGCCGTGGA | CCCCCTGATG | 1980 |
| 320 | CGCCGGCGG   | CCAAGACGGT  | GAACCTCCGC  | GTCCTCTACG | GCATGTCCGC | CCATAGGCTC | 2040 |
| 322 | TCCCAGGAGC  | TTGCCATCCC  | CTACGAGGAG  | GCGGTGGCCT | TTATAGAGGC | TACTTCCAAA | 2100 |
| 324 | GCTTCCCCAA  | GGTCGGGCC   | TGGATAGAAA  | AGACCTGGA  | GGAGGGGAGG | AAGCGGGCT  | 2160 |
| 326 | ACGTGGAAAC  | CCTCTTCGGA  | AGAAGGCGCT  | ACGTGCCCCG | CCTCAACGCC | CGGGTGAAGA | 2220 |
| 328 | GCGTCAGGGA  | GGCCGCGGAG  | CGCATGGCCT  | TCAACATGCC | CGTCCAGGGC | ACCGCCGCCG | 2280 |
| 330 | ACCTCATGAA  | GCTCGCCATG  | GTGAAGCTCT  | TCCCCCGCCT | CCGGGAGATG | GGGGCCCGCA | 2340 |
| 332 | TGCTCCTCCA  | GGTCCACGAC  | GAGCTCCTCC  | TGGAGGCC   | CCAAGCGCGG | GCCGAGGAGG | 2400 |
| 334 | TGGCGGCTTT  | GGCCAAGGAG  | GCCATGGAGA  | AGGCCTATCC | CCTCGCCGTG | CCCCTGGAGG | 2460 |
| 336 | TGGAGGTGGG  | GATGGGGAG   | GACTGGCTTT  | CCGCCAAGGG | TTAG       |            | 2504 |

338 (2) INFORMATION FOR SEQ ID NO: 4:

340 (i) SEQUENCE CHARACTERISTICS:

341 (A) LENGTH: 832 amino acids

342 (B) TYPE: amino acid

343 (C) STRANDEDNESS: single

344 (D) TOPOLOGY: linear

346 (ii) MOLECULE TYPE: protein

350 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

352 Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/08/520,946A

DATE: 06/21/2002  
TIME: 14:46:47

Input Set : A:\Seqsub2.app  
Output Set: N:\CRF3\06212002\H520946A.raw

|     |   |     |     |     |
|-----|---|-----|-----|-----|
| 353 | 1   | 5   | 10  | 15  |
| 355 | Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly |     |     |     |
| 356 | 20  | 25  | 30  |     |
| 358 | Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala |     |     |     |
| 359 | 35  | 40  | 45  |     |
| 361 | Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val |     |     |     |
| 362 | 50  | 55  | 60  |     |
| 364 | Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly |     |     |     |
| 365 | 65  | 70  | 75  | 80  |
| 367 | Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu |     |     |     |
| 368 | 85  | 90  | 95  |     |
| 370 | Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu |     |     |     |
| 371 | 100   | 105 | 110 |     |
| 373 | Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys |     |     |     |
| 374 | 115   | 120 | 125 |     |
| 376 | Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp |     |     |     |
| 377 | 130   | 135 | 140 |     |
| 379 | Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly |     |     |     |
| 380 | 145   | 150 | 155 | 160 |
| 382 | Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro |     |     |     |
| 383 | 165   | 170 | 175 |     |
| 385 | Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn |     |     |     |
| 386 | 180   | 185 | 190 |     |
| 388 | Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu |     |     |     |
| 389 | 195   | 200 | 205 |     |
| 391 | Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu |     |     |     |
| 392 | 210   | 215 | 220 |     |
| 394 | Lys Pro Ala Ile Arg Glu Ile Leu Ala His Met Asp Asp Leu Lys     |     |     |     |
| 395 | 225   | 230 | 235 | 240 |
| 397 | Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val |     |     |     |
| 398 | 245   | 250 | 255 |     |
| 400 | Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe |     |     |     |
| 401 | 260   | 265 | 270 |     |
| 403 | Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu |     |     |     |
| 404 | 275   | 280 | 285 |     |
| 406 | Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly |     |     |     |
| 407 | 290   | 295 | 300 |     |
| 409 | Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp |     |     |     |
| 410 | 305   | 310 | 315 | 320 |
| 412 | Leu Leu Ala Leu Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro     |     |     |     |
| 413 | 325   | 330 | 335 |     |
| 415 | Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu |     |     |     |
| 416 | 340   | 345 | 350 |     |
| 418 | Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro |     |     |     |
| 419 | 355   | 360 | 365 |     |
| 421 | Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn |     |     |     |
| 422 | 370   | 375 | 380 |     |
| 424 | Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu |     |     |     |
| 425 | 385   | 390 | 395 | 400 |

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/08/520,946A

DATE: 06/21/2002  
TIME: 14:46:48

Input Set : A:\Seqsub2.app  
Output Set: N:\CRF3\06212002\H520946A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 4,5,181,182,190,366,617,628,685,714,722,738,784,1022,1029  
Seq#:7; N Pos. 1038,1053,1098,1105,1206,1227,1244,1251,1252,1253,1350,1380  
Seq#:7; N Pos. 1497,1530,1569,1572,1641,1653,1655,1770,1812,2319,2346,2396  
Seq#:8; Xaa Pos.2,63,109,186,205,209,227,228,233,240,243,244,247,260,290  
Seq#:8; Xaa Pos.329,336,340,368,414,417,418,431,551,605,773,794,798,823,833

VERIFICATION SUMMARY  
PATENT APPLICATION: US/08/520,946A

DATE: 06/21/2002  
TIME: 14:46:48

Input Set : A:\Seqsub2.app  
Output Set: N:\CRF3\06212002\H520946A.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:1126 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0  
L:1135 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:48  
L:1144 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:96  
L:1159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:176  
L:1162 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:192  
L:1165 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:208  
L:1168 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:224  
L:1171 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:240  
L:1174 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:256  
L:1180 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:288  
L:1186 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:320  
L:1189 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:336  
L:1192 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:352  
L:1201 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:400  
L:1204 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:416  
L:1228 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:544  
L:1237 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:592  
L:1270 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:768  
L:1273 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:784  
L:1279 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:816  
L:1282 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:832